Method of Conjugate Radii for Solving Linear and Nonlinear Systems

Philip R. Nachtsheim

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the Lead Center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at *http://www.sti.nasa.gov*

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA Access Help Desk at (301) 621-0134

- Telephone the NASA Access Help Desk at (301) 621-0390

- Write to:
  NASA Access Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076-1320
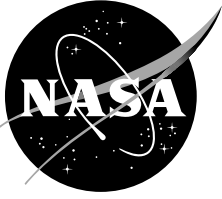
Method of Conjugate Radii for Solving Linear and Nonlinear Systems

Philip R. Nachtsheim
Ames Research Center, Moffett Field, California

Available from:

Method of Conjugate Radii for Solving Linear and Nonlinear Systems

PHILIP R. NACHTSHEIM

Ames Research Center

Summary

This paper describes a method to solve a system of N linear equations in N steps. A quadratic form is developed involving the sum of the squares of the residuals of the equations. Equating the quadratic form to a constant yields a surface which is an ellipsoid.

For different constants, a family of similar ellipsoids can be generated. Starting at an arbitrary point an orthogonal basis is constructed and the center of the family of similar ellipsoids is found in this basis by a sequence of projections. The coordinates of the center in this basis are the solution of the linear system of equations. A quadratic form in N variables requires N projections. That is, the current method is an exact method. It is shown that the sequence of projections is equivalent to a special case of the Gram-Schmidt orthogonalization process. The current method enjoys an advantage not shared by the classic Method of Conjugate Gradients. The current method can be extended to nonlinear systems without modification. For nonlinear equations the Method of Conjugate Gradients has to be augmented with a line-search procedure. Results for linear and nonlinear problems are presented.

Introduction

The computation method described herein belongs to the general class of methods where the iterates are derived from information available about the function and its derivatives at each step. It applies to functions where the minimum value of the function is known to be zero. In order to apply the method to the general multidimensional root finding problem, a function is introduced which is the sum of the squares of the individual functions. This function is called the squared-error function. It is positive definite and has a global minimum of zero at all solutions of the original set of equations.

Now consider the case of linear equations. In this case the sum of the squares of the residuals is considered. The squared-error function is a quadratic function of the N variables. This quadratic form is positive definite and has a global minimum of zero at all solutions of the original set of equations, provided that the rank of the

coefficient matrix is N. Equating the quadratic form to a constant yields a surface which is an ellipsoid. A basis for this space is formed from the set of vectors reciprocal to the conjugate radii of the ellipsoid. The conjugate radii and the reciprocal vectors are generated by an interlocking iterative procedure. This procedure is followed to solve a system of linear equations. The solution of N linear equations is accomplished in N iterations. That is, the current method is an exact method for linear systems.

An advantage of this method is that it can be extended to systems of nonlinear equations without modification. The Method of Conjugate Gradients (ref. 1) solves a linear system of dimension N in N steps. However, in order to extend this method to nonlinear equations a line search procedure has to augment the original method. The line search procedure is required at each stage to generate successive iteration points. At each stage the step size is the solution of a minimization problem along a line in a selected direction (ref. 2).

This paper has a twofold purpose. One purpose is to present the method of conjugate radii for linear systems. Another purpose is to extend the method to nonlinear systems. "Nonlinear problems can in most cases be solved only by approximating them by linear problems" (ref. 3). The approach of the current method to solve a nonlinear system is to approximate it by a quadratic function.

The current method is quadratically convergent. That is, for a quadratic function of N variables, the root will be found in N iterations. In the nonlinear case, assume the squared-error function can be expanded in a Taylor series about the root. The matrix whose elements are the second-order partial derivatives of the function is called the Hessian matrix. It is symmetric and positive definite. If it is assumed that the function can be approximated near the root by a quadratic form, the elements of the Hessian matrix are the coefficients of the quadratic form. For nonlinear functions, as the iterates approach the root, the squared-error function is more closely quadratic and hence convergence is more nearly quadratic.

After the relevant properties of central quadratic surfaces are identified, the method is described for linear systems and an algorithm is formulated. A linear and a nonlinear example are treated by the same algorithm. The relation

between the Gram-Schmidt orthogonalization procedure and the current method is discussed, as is the duality between the current method and the Method of Conjugate Gradients.

Method of Conjugate Radii

Problem Statement

The system of N linear equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ is solved by introducing the function

$$2f = (\mathbf{A} \cdot \mathbf{x} - \mathbf{b}) \cdot (\mathbf{A} \cdot \mathbf{x} - \mathbf{b}) = \mathbf{x} \cdot \mathbf{A}' \cdot \mathbf{A} \cdot \mathbf{x} - 2\mathbf{x} \cdot \mathbf{A}' \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{b} \tag{1}$$

The raised dot in an expression such as $\mathbf{A} \cdot \mathbf{x}$ denotes matrix multiplication. Square matrices appear as bold capital letters and vectors appear as bold lower case letters. In the dot product, vectors that appear before the dot are understood to be row vectors and those that appear after are understood to be column vectors. The transpose of a matrix is denoted by a prime.

For an arbitrary $\mathbf{x}$ the function (1) is the square of the magnitude of the residual $\mathbf{b} - \mathbf{A} \cdot \mathbf{x}$ which is a measure of the error incurred by the choice of $\mathbf{x}$. In carrying out the computations, it is necessary to evaluate the gradient.

The formula for the gradient is

$$\mathbf{g} \equiv \nabla f = \mathbf{A}' \cdot (\mathbf{A} \cdot \mathbf{x} - \mathbf{b}) \tag{2}$$

When the squared error is held constant, equation (1) represents a family of quadratic surfaces. Each constant value leads to a different surface. For a system of two linear equations the significance of the quadratic surfaces can be visualized. If the equations are expressed in terms of the usual Cartesian coordinates (x,y), and if the squared error represents altitude above the (x,y) plane, a three-dimensional surface is obtained. The surface has the form of a bowl, an elliptic paraboloid with vertex at the point $\mathbf{c}$ in the (x,y) plane. Figure 1 shows multiple level lines of such a surface. Two of these level lines are projected onto the (x,y) plane. The vector $\mathbf{c}$ is the solution of the system of equations, that is, $\mathbf{A} \cdot \mathbf{c} = \mathbf{b}$. The projections of the level lines of the squared-error function onto the (x,y) plane represents a family of concentric ellipses, quadratic surfaces, similar to each other. These are called error-ellipses (ref. 3). In higher dimensions, an error-ellipsoid is obtained.

The problem of solving the linear system $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ for $\mathbf{x}$ is equivalent to the problem of determining the center $\mathbf{c}$ of the family of similar error-ellipsoids generated by the squared-error function.

In view of the equivalence of the two problems, the following problem is addressed herein. The problem to be solved is: find the origin of a system of coordinates in which a family of similar central quadratic surfaces is embedded starting at an arbitrary point.

The central quadratic surfaces associated with the problem $\mathbf{A} \cdot \mathbf{r} = 0$ are related to the non-central quadratic surfaces associated with the problem $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ by means of the coordinate transformation $\mathbf{r} = \mathbf{x} - \mathbf{c}$. It is to be emphasized in what follows that $\mathbf{x}$ and $\mathbf{r}$ are the same point, and this point has different labels and components, in the two coordinate frames.

The point $\mathbf{x} = \mathbf{c}$ corresponds to the center of the family of similar error-ellipsoids in the $\mathbf{x}$ coordinate frame, and the point $\mathbf{r} = 0$ corresponds to the center of the same family in the $\mathbf{r}$ coordinate frame. It will be shown that the sequence of operations that leads to the origin in the $\mathbf{r}$ coordinate frame, starting from an arbitrary point $\mathbf{r}$, leads to the solution $\mathbf{c}$ in the $\mathbf{x}$ coordinate frame, starting from an arbitrary point $\mathbf{x}$.

The properties of central quadratic surfaces play a significant role in the analysis which will be presented. Denoting $\mathbf{A}' \cdot \mathbf{A}$ by $\mathbf{H}$, the expressions for the squared-error function and the gradient can be written in the $\mathbf{r}$ coordinate frame as follows:

$$2f = \mathbf{r} \cdot \mathbf{H} \cdot \mathbf{r} \tag{3}$$

$$\mathbf{g} = \mathbf{H} \cdot \mathbf{r} \tag{4}$$

The representation of planes in space and the essential properties of central quadratic surfaces which will be used in the analysis are identified. The three-dimensional case will be discussed.

Representation of Planes and Properties of Central Quadratic Surfaces

In analytic geometry the intercept form of a plane is written as

$$\mathbf{r} \cdot \mathbf{n} = 1$$

where the components of $\mathbf{n}$ are the reciprocals of the intercepts of the plane on the coordinate axes and $\mathbf{r}$ is the position vector of a general point on the plane.

This equation when rewritten as follows reveals some interesting features:

$$|\mathbf{r}| \cos(\mathbf{n}, \mathbf{r}) |\mathbf{n}| = 1$$

Here $|\mathbf{r}|$ and $|\mathbf{n}|$ denote the magnitudes of the vectors $\mathbf{r}$ and $\mathbf{n}$ respectively and $(\mathbf{n}, \mathbf{r})$ denotes the angle between $\mathbf{n}$ and $\mathbf{r}$. Now $|\mathbf{n}|$, the length of $\mathbf{n}$, is the reciprocal of the

perpendicular distance from the origin to the plane since $|\mathbf{r}|\cos(\mathbf{n},\mathbf{r})$ is that distance. As stated, the components of $\mathbf{n}$ are the reciprocals of the intercepts of the plane on the axes.

Gibbs (ref. 4) points out that the vector $\mathbf{n}$ can be used to denote the plane. It may be taken as the vector coordinate of the plane. To designate a plane denoted by the vector $\mathbf{n}$ the notation $(\mathbf{n},1)$ is used. This is the notation of analytic geometry to specify the "coordinates of a plane."

If the elements of $\mathbf{H}$ are constant, then

$$\mathbf{r}\cdot\mathbf{H}\cdot\mathbf{r} = \text{const}$$

is a quadratic in $\mathbf{r}$. It represents an ellipsoid with center at the origin.

Hence for the problem at hand, the following family of similar ellipsoids is considered:

$$\frac{\mathbf{r}\cdot\mathbf{H}\cdot\mathbf{r}}{2f} = 1 \qquad (5)$$

Define the normalized gradient

$$\mathbf{n} \equiv \frac{\mathbf{g}}{2f} = \frac{\mathbf{H}}{2f}\cdot\mathbf{r} \qquad (6)$$

From this definition it follows that $\mathbf{n}\cdot\mathbf{r} = 1$.

Figure 2 illustrates the foregoing relations in two dimensions. The tangent plane to an ellipse at the point R is shown. The normalized gradient at this point is also shown. This is the vector originating at the point R and terminating at the point N. The center of the ellipse is labeled C and this point is the origin of the $\mathbf{r}$ coordinate frame. The point labeled 0 in this figure is the origin of the $\mathbf{x}$ coordinate frame. The components of the vector $\mathbf{n}$, which is the normalized gradient referred to the origin in the $\mathbf{r}$ coordinate frame, are $(1,1)$. The intercepts of the tangent plane when referred to the same origin are $(1,1)$. The vector $\mathbf{n}$ represents the tangent plane. The length of this vector is $\sqrt{2}$. The distance of the tangent plane from the origin is $1/\sqrt{2}$ which is the reciprocal of the length of $\mathbf{n}$. Also, the vector which represents the plane tangent to the similar inner ellipse at the point $(0,.5)$ with intercepts $(.5,.5)$ would have components $(2,2)$ and length $2\cdot\sqrt{2}$. The distance of this plane from the origin is $1/(2\cdot\sqrt{2})$. That is, the nearer a plane is to the origin the longer the vector will be that represents it.

At any point $\mathbf{r}_1$ of the quadratic surface (5), $\mathbf{n}_1$, the vector given by equation (6) at that point, is normal to the tangent plane at $\mathbf{r}_1$. The equation of the tangent plane at $\mathbf{r}_1$ is $\mathbf{n}_1\cdot(\mathbf{r}-\mathbf{r}_1) = 0$. Since $\mathbf{r}_1\cdot\mathbf{n}_1 = 1$, this can be written $\mathbf{n}_1\cdot\mathbf{r} = 1$. Now $|\mathbf{n}_1|$ the length of $\mathbf{n}_1$ is the

reciprocal of the perpendicular distance from the origin to the tangent plane at $\mathbf{r}_1$. The plane coordinates of the tangent plane are $(\mathbf{n}_1, 1)$.

The vector $\dfrac{\mathbf{n}_1}{\mathbf{n}_1\cdot\mathbf{n}_1}$ is the vector drawn from the center, the origin of the $\mathbf{r}$ coordinate frame, perpendicular to the tangent plane (ref. 4). See figure 2. The standard approach for solving a system of equations is to assume an initial approximation $\mathbf{r}_1$ and to proceed to an improved approximation $\mathbf{r}_2$ by using an iterative formula of the form

$$\mathbf{r}_2 = \mathbf{r}_1 - \frac{\mathbf{n}_1}{\mathbf{n}_1\cdot\mathbf{n}_1}$$

The significance of the vector drawn from the origin perpendicular to the tangent plane is that it represents, in the terminology of the standard approach, the negative of the "direction of search," and its magnitude determines the "step size" in this direction. As explained below, the strategy of establishing a plane at a point $\mathbf{r}_1$ and moving toward the origin a distance $\dfrac{1}{|\mathbf{n}_1|}$ is the basis of the current method.

Any two vectors $\mathbf{r}$ and $\mathbf{s}$ are referred to as conjugate vectors if $\mathbf{r}\cdot\mathbf{H}\cdot\mathbf{s} = 0$. The tangent plane at the point $\mathbf{r}_1$ is $(\mathbf{n}_1, 1)$. The plane through the origin parallel to this tangent plane has plane coordinates $(\mathbf{n}_1, 0)$. It is referred to as the diametral plane conjugate with $\mathbf{r}_1$. Any vector $\mathbf{r}$ in that plane satisfies the relation $\mathbf{r}\cdot\mathbf{H}\cdot\mathbf{r}_1 = \mathbf{r}\cdot\mathbf{n}_1 = 0$.

If three points located on similar ellipsoids of the form (5) satisfy the equations

$$\mathbf{r}_1\cdot\mathbf{H}\cdot\mathbf{r}_2 = \mathbf{r}_2\cdot\mathbf{H}\cdot\mathbf{r}_3 = \mathbf{r}_3\cdot\mathbf{H}\cdot\mathbf{r}_1 = 0$$

the position vectors $(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ are said to form a conjugate set. The vectors $(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ and $(\mathbf{n}_1,\mathbf{n}_2,\mathbf{n}_3)$ form reciprocal sets. See reference 4. The relation $\mathbf{r}_1\cdot\mathbf{H}\cdot\mathbf{r}_2$ is symmetric in the indices 1 and 2; that is, $\mathbf{r}_1\cdot\mathbf{H}\cdot\mathbf{r}_2 = \mathbf{r}_2\cdot\mathbf{H}\cdot\mathbf{r}_1$. Hence, if $\mathbf{n}_1\cdot\mathbf{r}_2 = 0$ then $\mathbf{n}_2\cdot\mathbf{r}_1 = 0$.

Method of Solution

The method of solution is similar to the process of assigning Cartesian coordinates (x,y,z), given the Cartesian frame, to a point P in space. The frame is given by giving three mutually perpendicular coordinate planes. Their lines of intersection are the axes of the coordinates

3

and are called the x axis, y axis, and z axis. The point of intersection of the coordinate planes is the origin. Let P be any point in space and let three planes be drawn through P perpendicular to the coordinate planes and cutting the axes at x, y, and z. These three planes together with the coordinate planes bound a rectangular parallelepiped, of which P and the origin 0 are opposite vertices. The three edges x, y, and z are called the rectangular coordinates of P. That is, the rectangular coordinates of P are equal to its perpendicular distances from the coordinate planes.

The method of solution described herein follows the procedure described in the previous paragraph starting at the point P. That is, the origin is located with respect to P. The length of the edges of the parallelepiped are conveniently found using the vector representation of a plane. The strategy at each iteration point for the construction of a plane is enabled by the relations satisfied by the mutually conjugate position vectors generated up to that point.

The construction will be carried out for the three-dimensional case. Figure 3 shows the parallelepiped. The vertex of the parallelepiped opposite the origin 0 is labeled R_1. The vertex of the parallelogram in the plane of the three points 0, R_2, and R_3 opposite the origin 0 is labeled R_2. The vertex of the line connecting the points 0 and R_3 opposite the origin 0 is labeled R_3.

In three dimensions, three iteration points $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ will be generated. It will be demonstrated that the iteration points form a conjugate set. In carrying out the procedure the following results will be demonstrated:

1) The vectors $\mathbf{n}_1, \mathbf{n}_1 + \mathbf{n}_2, \mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3$ form a three-dimensional orthogonal basis.

2) The vectors reciprocal to the vectors $\mathbf{n}_1, \mathbf{n}_1 + \mathbf{n}_2, \mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3$ satisfy the following relations

$$\mathbf{r}_1 - \mathbf{r}_2 = \frac{\mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_1} \tag{7}$$

$$\mathbf{r}_2 - \mathbf{r}_3 = \frac{\mathbf{n}_1 + \mathbf{n}_2}{(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2)} \tag{8}$$

$$\mathbf{r}_3 = \frac{\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3}{(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) \cdot (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)} \tag{9}$$

The reciprocal vectors are the right hand members of equations (7), (8), and (9) respectively.

Carrying out the steps of the procedure will be a constructive demonstration of the above results. The

procedure can be regarded as a coordinate transformation from a given basis to a new basis. The coordinate transformation will be accomplished by a sequence of projections orthogonal to each of the vectors $\mathbf{n}_1, \mathbf{n}_1 + \mathbf{n}_2, \mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3$. These vectors themselves are obtained during the sequence.

The problem of finding the origin will be carried out in the new basis. Referring to figure 3, R_1 will be projected onto the parallelogram yielding R_2. Then R_2 will be projected onto the line $0 - R_3$ yielding R_3. Finally, R_3 will be projected onto the origin. These projections will be identified at each step.

At each step of the iteration, it is necessary to compute the components of the original basis $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ in terms of the new basis $\mathbf{n}_1, \mathbf{n}_1 + \mathbf{n}_2, \mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3$. These components are needed to carry out the next step of the iteration or to show that the iteration is complete.

The process starts at an arbitrary point. In general, once $\mathbf{r}$ is known, $\mathbf{n}$ can be calculated using (6). At the point $\mathbf{r}_1$, construct the plane ($\mathbf{n}_1$, 1). That is, use the point normal form of the equation of a plane to write $\mathbf{n}_1 \cdot (\mathbf{r} - \mathbf{r}_1) = 0$. Since $\mathbf{n}_1 \cdot \mathbf{r}_1 = 1$, the intercept form of the plane is $\mathbf{n}_1 \cdot \mathbf{r} = 1$. This plane is the tangent plane at the point $\mathbf{r}_1$. Then construct a plane parallel to the plane at $\mathbf{r}_1$ that passes through the origin. This is the plane ($\mathbf{n}_1$, 0). This is the first coordinate plane. The distance of the original plane ($\mathbf{n}_1$, 1) from the origin is $\frac{1}{|\mathbf{n}_1|}$. The location of a point $\mathbf{r}_2$ that lies on the Cartesian coordinate plane ($\mathbf{n}_1$, 0), the parallelogram in figure 3, is obtained by projecting $\mathbf{r}_1$ onto that plane. That is

$$\mathbf{r}_2 = \mathbf{r}_1 - \frac{\mathbf{n}_1 \mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_1} \cdot \mathbf{r}_1$$

However, $\mathbf{r}_1$ lies on the tangent plane $(\mathbf{n}_1, 1)$; therefore, $\mathbf{n}_1 \cdot \mathbf{r}_1 = 1$. Hence, the location of the point is given by the equation

$$\mathbf{r}_2 = \mathbf{r}_1 - \frac{\mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_1} \tag{10}$$

The scalar product $\mathbf{n}_1 \cdot \mathbf{r}_2 = \mathbf{r}_1 \cdot \mathbf{H} \cdot \mathbf{r}_2 = 0$ shows that the two position vectors $\mathbf{r}_1$, $\mathbf{r}_2$ form a conjugate set since $\mathbf{n}_1 \cdot \mathbf{r}_2 = 0$. By symmetry $\mathbf{n}_2 \cdot \mathbf{r}_1 = 0$. That is, $\mathbf{n}_1 \cdot \mathbf{r}_2 = \mathbf{r}_1 \cdot \mathbf{H} \cdot \mathbf{r}_2 = \mathbf{r}_1 \cdot \mathbf{n}_2$.

4

Hence, the vectors $(\mathbf{r}_1, \mathbf{r}_2)$ and $(\mathbf{n}_1, \mathbf{n}_2)$ form reciprocal sets.

Therefore

$$(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) = 0 \qquad (11)$$

Equations (7) and (11) show that

$$\mathbf{n}_1 \cdot (\mathbf{n}_1 + \mathbf{n}_2) = 0 \qquad (12)$$

The following relation follows from equation (12):

$$\mathbf{n}_2 \cdot (\mathbf{n}_1 + \mathbf{n}_2) = (\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2) \qquad (13)$$

The relations (12) and (13) are needed to express the components of the basis $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ in terms of the new basis $\mathbf{n}_1, \mathbf{n}_1 + \mathbf{n}_2, \mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3$. These relations for two dimensions are summarized in the following multiplication table:

$\cdot$	$\mathbf{r}_1 - \mathbf{r}_2$ $\dfrac{\mathbf{n}_1}{\|\mathbf{n}_1\|^2}$	$\dfrac{\mathbf{n}_1 + \mathbf{n}_2}{\|\mathbf{n}_1 + \mathbf{n}_2\|^2}$
$\mathbf{n}_1$	1	0
$\mathbf{n}_2$	-1	1

In the table an entry in the dashed row is equal to the entry in the corresponding column in the row directly below it.

The projection of $\mathbf{n}_2$ onto the $(\mathbf{n}_1, 0)$ plane is $\mathbf{n}_1 + \mathbf{n}_2$. That is

$$\mathbf{n}_1 + \mathbf{n}_2 = \mathbf{n}_2 - \frac{\mathbf{n}_1 \mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_1} \cdot \mathbf{n}_2 \qquad (14)$$

The multiplication table was used to arrive at this equation.

Equation (12) suggests the vector $(\mathbf{n}_1 + \mathbf{n}_2)$ can be used as the normal to construct a plane perpendicular to the plane $(\mathbf{n}_1, 0)$. At the point $\mathbf{r}_2$, construct the plane $(\mathbf{n}_1 + \mathbf{n}_2, 1)$. The plane parallel to this plane that passes through the origin is $(\mathbf{n}_1 + \mathbf{n}_2, 0)$. Since $(\mathbf{n}_1, 0)$ is the first coordinate plane, it is appropriate to select $(\mathbf{n}_1 + \mathbf{n}_2, 0)$ as the second coordinate plane.

The location of a point $\mathbf{r}_3$ that lies on the second coordinate plane, $(\mathbf{n}_1 + \mathbf{n}_2, 0)$, is obtained by projecting $\mathbf{r}_2$ onto that plane. That is

$$\mathbf{r}_3 = \mathbf{r}_2 - \frac{(\mathbf{n}_1 + \mathbf{n}_2)(\mathbf{n}_1 + \mathbf{n}_2)}{(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2)} \cdot \mathbf{r}_2$$

However, $\mathbf{r}_2$ lies on the plane $(\mathbf{n}_1 + \mathbf{n}_2, 1)$; therefore, $(\mathbf{n}_1 + \mathbf{n}_2) \cdot \mathbf{r}_2 = 1$.

Hence, the location of the point is given by the equation

$$\mathbf{r}_3 = \mathbf{r}_2 - \frac{\mathbf{n}_1 + \mathbf{n}_2}{(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2)} \qquad (15)$$

Using equation (15) to represent $\mathbf{r}_3$ and utilizing the relations (12) and (13) which are summarized in the multiplication table, it is readily verified that $\mathbf{n}_1 \cdot \mathbf{r}_3 = \mathbf{n}_2 \cdot \mathbf{r}_3 = 0$ and by symmetry $\mathbf{n}_3 \cdot \mathbf{r}_1 = \mathbf{n}_3 \cdot \mathbf{r}_2 = 0$.

The geometric interpretation of this result is interesting. The equation $\mathbf{n}_1 \cdot \mathbf{r} = 0$ is the equation of the diametral plane conjugate with $\mathbf{r}_1$, and the equation $\mathbf{n}_2 \cdot \mathbf{r} = 0$ is the equation of the diametral plane conjugate with $\mathbf{r}_2$. Simultaneously, these two equations represent the line of intersection of the two diametral planes. The results, $\mathbf{n}_1 \cdot \mathbf{r}_3 = \mathbf{n}_2 \cdot \mathbf{r}_3 = 0$, establish that $\mathbf{r}_3$ is along this line of intersection. This is the line 0–R_3 shown in figure 3.

The conjugate relation is symmetric. Hence, $\mathbf{r}_1$ is in the plane conjugate with $\mathbf{r}_3$ as is $\mathbf{r}_2$. That is, $\mathbf{r}_2 \cdot \mathbf{H} \cdot \mathbf{r}_3 = \mathbf{r}_3 \cdot \mathbf{H} \cdot \mathbf{r}_1 = 0$. This result along with the result $\mathbf{r}_1 \cdot \mathbf{H} \cdot \mathbf{r}_2 = 0$ established previously indicates that the position vectors $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ form a conjugate set.

Since the vectors $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ and $(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3)$ form reciprocal sets, it follows

$$(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) \cdot (\mathbf{r}_2 - \mathbf{r}_1) = 0 \qquad (16)$$

$$(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) \cdot (\mathbf{r}_3 - \mathbf{r}_2) = 0 \qquad (17)$$

Equations (7) and (16) show

$$\mathbf{n}_1 \cdot (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) = 0 \qquad (18)$$

Equations (8) and (17) show

$$(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) = 0 \qquad (19)$$

From equations (18) and (19) it follows

$$\mathbf{n}_2 \cdot (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) = 0 \qquad (20)$$

The final relation needed to express the components of the basis $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ in terms of the new basis $\mathbf{n}_1, \mathbf{n}_1 + \mathbf{n}_2, \mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3$ can be obtained by combining equations (18) and (20). This relation is

$$\mathbf{n}_3 \cdot (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) = (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) \cdot (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) \qquad (21)$$

The foregoing relations for three dimensions are used to update the two-dimensional multiplication table. The current relations are summarized in the following multiplication table:

$\cdot$	$\mathbf{r}_1 - \mathbf{r}_2$ $\dfrac{\mathbf{n}_1}{\lvert \mathbf{n}_1 \rvert^2}$	$\mathbf{r}_2 - \mathbf{r}_3$ $\dfrac{\mathbf{n}_1 + \mathbf{n}_2}{\lvert \mathbf{n}_1 + \mathbf{n}_2 \rvert^2}$	$\dfrac{\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3}{\lvert \mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3 \rvert^2}$
$\mathbf{n}_1$	1	0	0
$\mathbf{n}_2$	-1	1	0
$\mathbf{n}_3$	0	-1	1

The projection of $\mathbf{n}_3$ onto the plane ($\mathbf{n}_1 + \mathbf{n}_2$, 0) is $\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3$. That is

$$\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3 = \mathbf{n}_3 - \frac{(\mathbf{n}_1 + \mathbf{n}_2)(\mathbf{n}_1 + \mathbf{n}_2)}{(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2)} \cdot \mathbf{n}_3 \tag{22}$$

The current multiplication table was used to arrive at this equation.

Equation (19) suggests the vector $(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)$ can be used as the normal to construct a plane perpendicular to the plane ($\mathbf{n}_1 + \mathbf{n}_2$, 0). At the point $\mathbf{r}_3$, construct the plane $(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3, 1)$. The plane parallel to this plane that passes through the origin is $(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3, 0)$. Equation (18) shows that this plane is perpendicular to the plane ($\mathbf{n}_1$, 0), and equation (19) shows that it is perpendicular to the plane ($\mathbf{n}_1 + \mathbf{n}_2$, 0). Since ($\mathbf{n}_1 + \mathbf{n}_2$, 0) is the second coordinate plane, it is appropriate to select $(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3, 0)$ as the third and final coordinate plane.

The location of a point $\mathbf{r}_4$ that lies on the third coordinate plane ($\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3$, 0) is obtained by projecting $\mathbf{r}_3$ onto that plane. That is

$$\mathbf{r}_4 = \mathbf{r}_3 - \frac{(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)}{(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) \cdot (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)} \cdot \mathbf{r}_3$$

However, $\mathbf{r}_3$ lies on the plane $(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3, 1)$; therefore, $(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) \cdot \mathbf{r}_3 = 1$. Hence, the location of the point is given by the equation:

$$\mathbf{r}_4 = \mathbf{r}_3 - \frac{\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3}{(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) \cdot (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)} \tag{23}$$

This equation will be used to show $\mathbf{r}_4 = 0$.

In order to express a vector in a basis $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$, its components in the reciprocal basis $(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3)$ are formed. Hence $\mathbf{r}_4$ can be expressed as follows:

$$\mathbf{r}_4 = (\mathbf{r}_4 \cdot \mathbf{n}_1)\mathbf{r}_1 + (\mathbf{r}_4 \cdot \mathbf{n}_2)\mathbf{r}_2 + (\mathbf{r}_4 \cdot \mathbf{n}_3)\mathbf{r}_3 \tag{24}$$

Using equation (23) to represent $\mathbf{r}_4$, the components can be calculated using equations (17), (19), and (21) which are summarized in the current multiplication table. All the components vanish, hence $\mathbf{r}_4 = 0$.

The above procedure, carried out in the $\mathbf{r}$ coordinate frame, led to the center, $\mathbf{r} = 0$, starting at an arbitrary point. The same procedure, carried out in the $\mathbf{x}$ coordinate frame, will also lead to the center, $\mathbf{x} = \mathbf{c}$, starting at an arbitrary point. That is, $\mathbf{r}_4 = 0$ corresponds to $\mathbf{x}_4 = \mathbf{c}$.

Algorithm

For any two points, 1 and 2 for instance, it is the case that $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{r}_1 - \mathbf{r}_2$. This equation is correct since it is independent of the origin. Using this fact and expressing the terms involving $\mathbf{r}$ in equations (7), (8), and (9) in terms of $\mathbf{x}$ and rearranging those equations leads to the following procedure in the $\mathbf{x}$ coordinate frame.

The procedure in the $\mathbf{x}$ coordinate frame starting from an arbitrary point, $\mathbf{x}_1$, can be summarized as follows:

$$\mathbf{x}_2 = \mathbf{x}_1 - \frac{\mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_1}$$

$$\mathbf{x}_3 = \mathbf{x}_2 - \frac{\mathbf{n}_1 + \mathbf{n}_2}{(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2)}$$

$$\mathbf{c} = \mathbf{x}_3 - \frac{\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3}{(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3) \cdot (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)}$$

Given $\mathbf{x}_1$ the above equations can be evaluated in sequence using the relation $\mathbf{n} = \mathbf{g}/2f$. In general, once $\mathbf{x}$ is known, $\mathbf{g}$ and 2f can be calculated using equations (1) and (2) for the linear case. For the linear case, it is understood that $\mathbf{A}$ is an NxN matrix of full rank. For the nonlinear case the gradient vector $\mathbf{g} = \nabla f$ is obtained by calculating the required partial derivatives of the squared-error-function 2f.

For N dimensions the algorithm is implemented by carrying out the following steps:

Set $\mathbf{n} = 0$

Set $\mathbf{x} = 0$ or any arbitrary value

for index = 1:1:N where N is the number of unknowns

$\mathbf{n} = \mathbf{n} + \mathbf{g}/2f$

$$\mathbf{x} = \mathbf{x} - \frac{\mathbf{n}}{\mathbf{n} \cdot \mathbf{n}}$$

end

return $\mathbf{x}$

Examples

A Linear Example

A two-dimensional example will be presented. This example will illustrate the application of the method and will also illustrate the relation between the basis $(\mathbf{r}_1, \mathbf{r}_2)$ and its reciprocal $(\mathbf{n}_1, \mathbf{n}_2)$ and the relation between the orthogonal basis $(\mathbf{n}_1, \mathbf{n}_1 + \mathbf{n}_2)$ and its reciprocal $(\dfrac{\mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_2}, \dfrac{\mathbf{n}_1 + \mathbf{n}_2}{(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2)})$. The solution of the following linear system is considered:

$$3x + y = 5$$
$$x + 2y = 5$$

Equations (1) and (2) are used to compute 2f and $\mathbf{g}$.

Figure 4 shows the projection of two level lines of the squared-error function, 2f, onto the (x,y) plane. The solution of the system of equations is (x,y) = (1,2). This point is labeled C in the figure. The origin of the $\mathbf{x}$ coordinate frame is labeled 0. The starting point was (x,y) = (1,3). The two vectors forming the basis $(\mathbf{r}_1, \mathbf{r}_2)$ are labeled (R_1, R_2) in the figure. The vectors $(\mathbf{n}_1, \mathbf{n}_2)$ reciprocal to them are also shown and are labeled (N_1, N_2). In addition, the parallelogram of which R_1 is a diagonal is shown. The vertices of this parallelogram are labeled C, P, R_1, and R_2. The portion of the parallelogram that is not obscured by solid lines is shown as the dashed line. The line R_1–P is tangent to the ellipse at the point R_1. The vector $\mathbf{n}_1 + \mathbf{n}_2$ terminates at the point Q. The magnitude of both of the vectors $\mathbf{n}_1$ and $\mathbf{n}_1 + \mathbf{n}_2$ is $\sqrt{2}$.

The vectors $(\dfrac{\mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_2}, \dfrac{\mathbf{n}_1 + \mathbf{n}_2}{(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2)})$ reciprocal to the base vectors $(\mathbf{n}_1, \mathbf{n}_1 + \mathbf{n}_2)$ are not labeled in the figure. They are in the same direction as the base vectors

and their magnitudes are the reciprocals of the base vectors, namely $1/\sqrt{2}$. The first reciprocal vector terminates at the point P and the second terminates at the point R_2 . They are the edges of the parallelogram of which R_1 is the diagonal.

In order to describe the sequence of projections, the same problem as above is discussed but with a different starting point. Figure 5 identifies the projection planes. All the labels shown in this figure have the same significance as the labels in figure 4. The starting point used in figure 5 was (x,y) = (2,0). The identification of the four planes used in the construction is as follows: The $(\mathbf{n}_1, 1)$ plane is the horizontal plane that passes through the point R_1 . The parallel horizontal plane $(\mathbf{n}_1, 0)$ is the plane that passes through the point C, the center. The $(\mathbf{n}_1 + \mathbf{n}_2, 1)$ plane is the vertical plane that passes through the point R_2 . The parallel vertical plane $(\mathbf{n}_1 + \mathbf{n}_2, 0)$ is the plane that passes through the point C.

The projections described earlier were performed. First, $\mathbf{r}_1$ which is located in horizontal plane $(\mathbf{n}_1, 1)$ was projected onto the parallel $(\mathbf{n}_1, 0)$ plane, as was $\mathbf{n}_2$. This led to the position vector $\mathbf{r}_2$ and the vector $\mathbf{n}_1 + \mathbf{n}_2$ respectively. The vectors $\mathbf{r}_2$ and $\mathbf{n}_1 + \mathbf{n}_2$ both terminate at the point R_2 . Second, $\mathbf{r}_2$, which is located on the vertical plane $(\mathbf{n}_1 + \mathbf{n}_2, 1)$, was projected onto the parallel $(\mathbf{n}_1 + \mathbf{n}_2, 0)$ plane. This results in the point C which is the solution of the system of equations. The orthogonal planes which were constructed are $(\mathbf{n}_1, 0)$ and $(\mathbf{n}_1 + \mathbf{n}_2, 0)$. The vectors $(\dfrac{\mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_2}, \dfrac{\mathbf{n}_1 + \mathbf{n}_2}{(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2)})$ which terminate at the points (P, R_2) are the edges of the parallelogram of which R_1 is the diagonal.

The results shown in figure 5 permit a comparison to be made between the current method and the Method of Conjugate Gradients. If the calculation was performed by that method starting at the point R_1 , the next point would be the point labeled CG in the figure. This point is located by minimizing the squared-error function along the line $R_1 - R_2$ shown in the figure. The radii C-CG and C-P are conjugate radii.

A Nonlinear Example

The extension of the current method to nonlinear systems will be illustrated for Rosenbrock's function (ref. 5).

$$2f(x, y) = 100 (y - x^2)^2 + (1 - x)^2$$

The current method can be applied to find the root $(x, y) = (1,1)$, where $f(x, y) = 0$, since Rosenbrock's function represents the sum of the squares of two individual functions. The gradient vector $\mathbf{g} = \nabla f$ is readily found by calculating the required partial derivatives of the function f. A contour map is shown in figure 6 which shows the solution path to the root for the current method starting at the point (–1.2,1). The function value was reduced to 4×10^{-13} in 13 iterations. This compares with Fletcher and Reeves' value of 1×10^{-8} after 27 iterations using a line search procedure together with the Method of Conjugate Gradients (ref. 2).

Another contour map is shown in figure 7 which shows the solution path to the root for the current method starting at the point (–1.92,2). Results for this starting point are given in reference 6. These results are compared with the current method in the table below. For these results the function value was less than 1×10^{-4}.

Method	Iterations for convergence
Gauss-Newton	48
Levenberg-Marquardt	90
Current	10

Discussion of Results

The sequence of projections orthogonal to each of the vectors $\mathbf{n}_1, \mathbf{n}_1 + \mathbf{n}_2, \mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3$ is identical to the Gram-Schmidt orthogonalization process applied to the vectors $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$. The results of the Gram-Schmidt process are shown below:

$$\mathbf{n}_1 = \mathbf{n}_1$$

$$\mathbf{n}_1 + \mathbf{n}_2 = \mathbf{n}_2 - \frac{\mathbf{n}_1 \mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_1} \cdot \mathbf{n}_2 \qquad (25)$$

$$\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3 = \mathbf{n}_3 - \frac{\mathbf{n}_1 \mathbf{n}_1}{\mathbf{n}_1 \cdot \mathbf{n}_1} \cdot \mathbf{n}_3 - \frac{(\mathbf{n}_1 + \mathbf{n}_2)(\mathbf{n}_1 + \mathbf{n}_2)}{(\mathbf{n}_1 + \mathbf{n}_2) \cdot (\mathbf{n}_1 + \mathbf{n}_2)} \cdot \mathbf{n}_3 \qquad (26)$$

Equations (12) and (18) combined show that $\mathbf{n}_1 \cdot \mathbf{n}_3 = 0$. The same result is obtained by consulting the current multiplication table. It can be seen that there is no need to perform the first subtraction in equation (26) since the term to be subtracted is zero. Hence, the sequence of projections is identical to the Gram-Schmidt orthogonalization process applied to the vectors $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$.

That they are identical for all cases can be seen by rewriting equations (25) and (26) employing equations (7) and (8) which give alternate expressions for the vectors reciprocal to $\mathbf{n}_1$ and $\mathbf{n}_1 + \mathbf{n}_2$.

The Gram-Schmidt equations (25) and (26), rewritten employing equations (7) and (8), are shown below:

$$\mathbf{n}_1 + \mathbf{n}_2 = \mathbf{n}_2 - \mathbf{n}_1(\mathbf{r}_1 - \mathbf{r}_2) \cdot \mathbf{n}_2 \qquad (27)$$

$$\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3 = \mathbf{n}_3 - \mathbf{n}_1(\mathbf{r}_1 - \mathbf{r}_2) \cdot \mathbf{n}_3 - (\mathbf{n}_1 + \mathbf{n}_2)(\mathbf{r}_2 - \mathbf{r}_3) \cdot \mathbf{n}_3 \qquad (28)$$

The scalar products involving the vectors of the reciprocal sets $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ and $(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3)$ which are needed to evaluate the terms in equations (27) and (28) are arranged in a table below:

$(\mathbf{r}_1 - \mathbf{r}_2) \cdot \mathbf{n}_2$	
$(\mathbf{r}_1 - \mathbf{r}_2) \cdot \mathbf{n}_3$	$(\mathbf{r}_2 - \mathbf{r}_3) \cdot \mathbf{n}_3$

The only non-vanishing scalar products lie along the diagonal. For a given row, the terms that appear before the term on the diagonal all involve vanishing scalar products of the reciprocal sets of vectors. Hence, only one subtraction per step is required to carry out the Gram-Schmidt process, and the term subtracted at each step is identical to the term subtracted in the projection process. The same result is obtained directly using equations (25) and (26) if the current multiplication table is consulted to determine the value of the scalar products.

It should be noted that the Gram-Schmidt procedure can be used to derive the basic algorithm of the Method of Conjugate Gradients (ref. 7). As is the case here, in that derivation only one subtraction per step is required.

There is a relatively direct duality between the Method of Conjugate Gradients and the current method. In all approaches to root finding, starting at an arbitrary point, the following requirements have to be addressed: In proceeding from one iteration point to the next, a step size and a direction are required. In the Method of Conjugate Gradients, consecutive directions are in conjugate directions. The step size is such that the gradients at consecutive points are orthogonal. In the current method, consecutive directions are orthogonal. The step size is such that consecutive points lie on conjugate radii.

For the three-dimensional case, it is of interest to note that the calculation performed herein parallels a similar calculation performed in x-ray diffraction studies. The vectors in the **r** coordinate frame correspond to lattice vectors, and the vectors in the **n** coordinate frame correspond to reciprocal lattice vectors. In x-ray diffraction studies, the purpose of the calculation is to determine the spacing between lattice planes. This interplanar spacing is used to determine the condition for scattering-in-phase known as Bragg's law.

For systems of nonlinear equations, the current method is a general method that converges quadratically to the solution. That is, for a quadratic function of N variables, it converges to the solution in N steps. The current method enables a nonlinear system to be approximated by a quadratic function.

References

1. Hestenes, M.; and Stiefel, E.: *Method of Conjugate Gradients for Solving Linear Systems.* Report 1659, Nat. Bur. Standards, 1952.

2. Fletcher, R.; and Reeves, C. M.: Function Minimization by Conjugate Gradients. *The Computer Journal*, vol. 7, 1964, pp. 149–154.

3. Stiefel, E.: *An Introduction to Numerical Mathematics*. Academic Press, 1969.

4. Gibbs-Wilson: *Vector Analysis.* Yale University Press, 1901.

5. Rosenbrock, H. H.: An Automatic Method for Finding Stationary Values of a Function of Several Variables. *The Computer Journal*, vol. 3, 1960, p. 175.

6. Grace, A.: Optimization TOOLBOX for use with MATLAB™, Nov. 1992. (MATLAB™ is a registered trademark of The MathWorks, Inc., 24 Prime Park Way, Natick, MA.)

7. Beckman, F. S.: The Solution of Linear Equations by the Conjugate Gradient Method. In *Mathematical Methods for Digital Computers.* A. Ralston and H. S. Wilf, eds., Wiley, 1960.
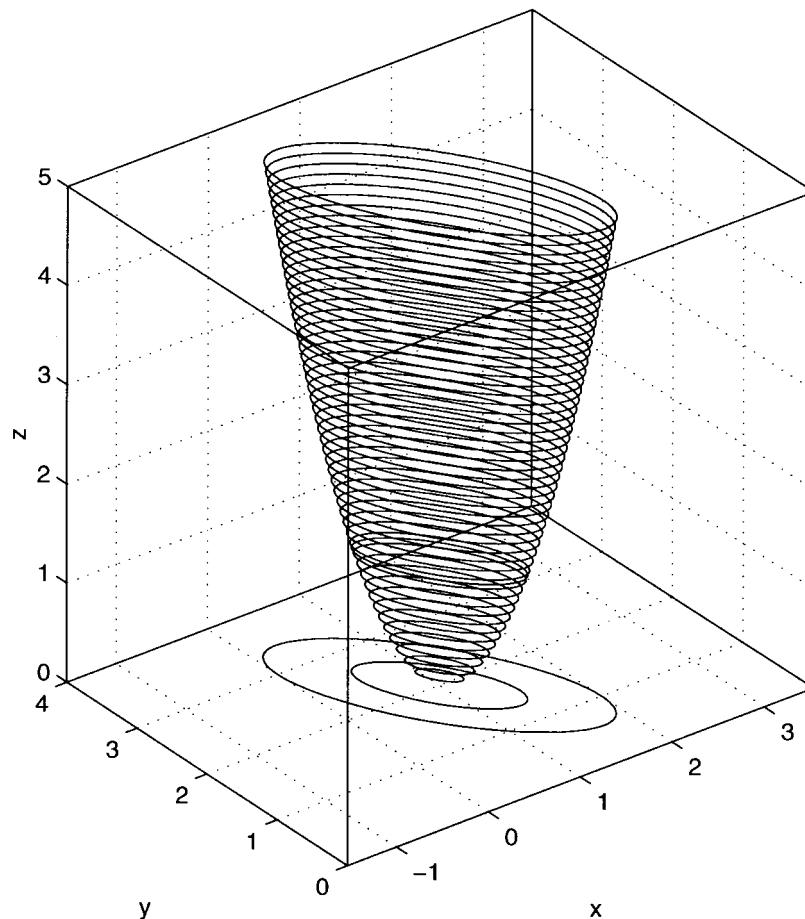
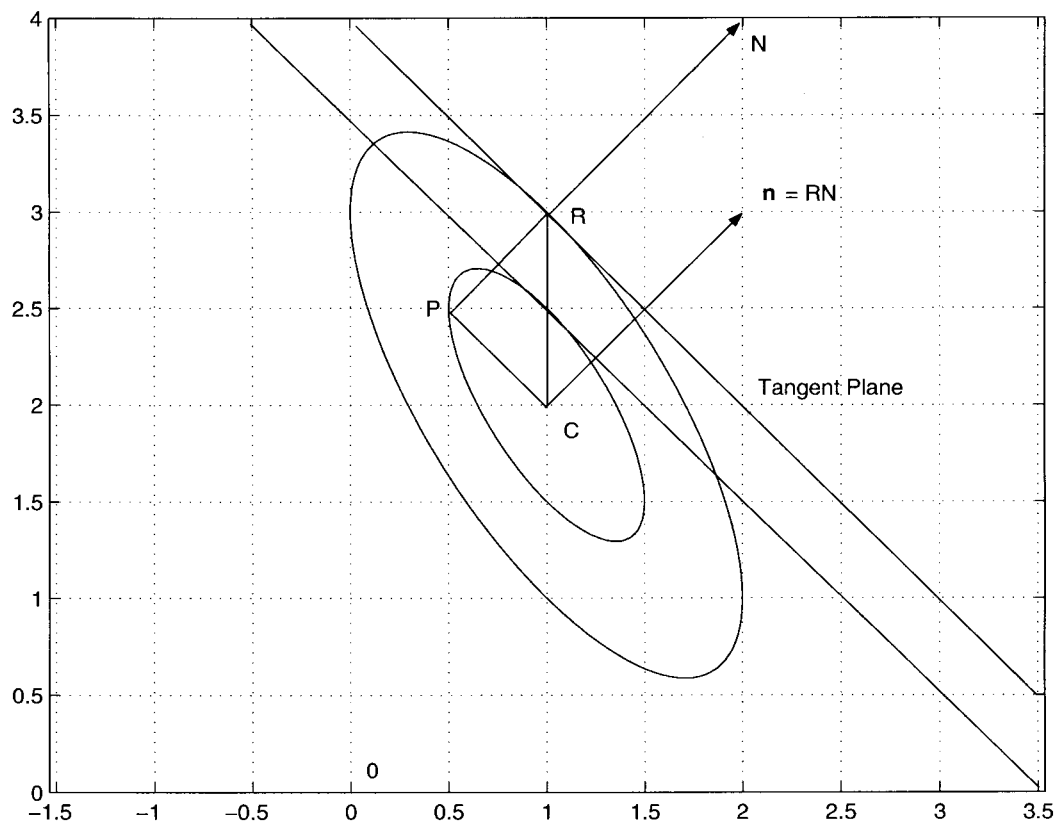*Figure 1. The squared-error function.*
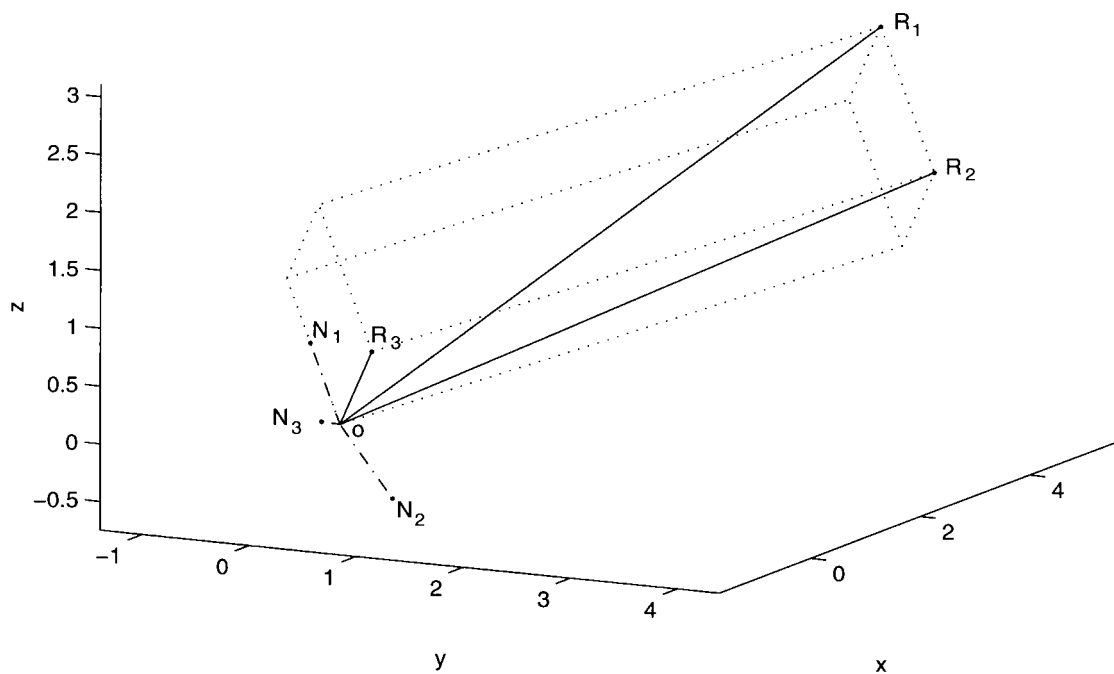
Figure 2. Vector representation of a plane.

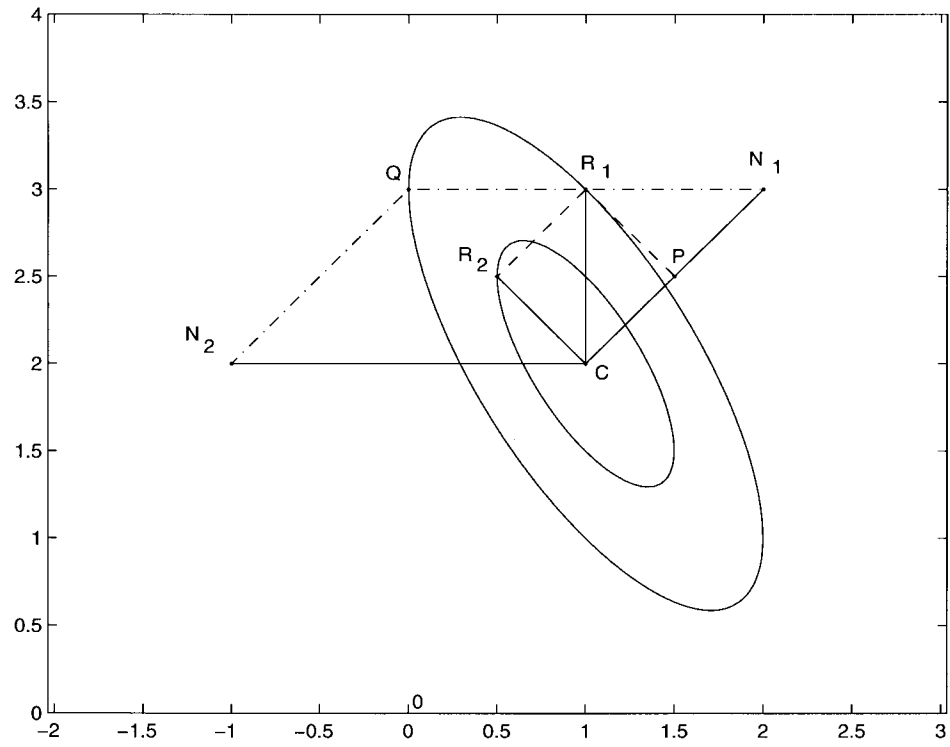Figure 3. The conjugate set and the reciprocal set.
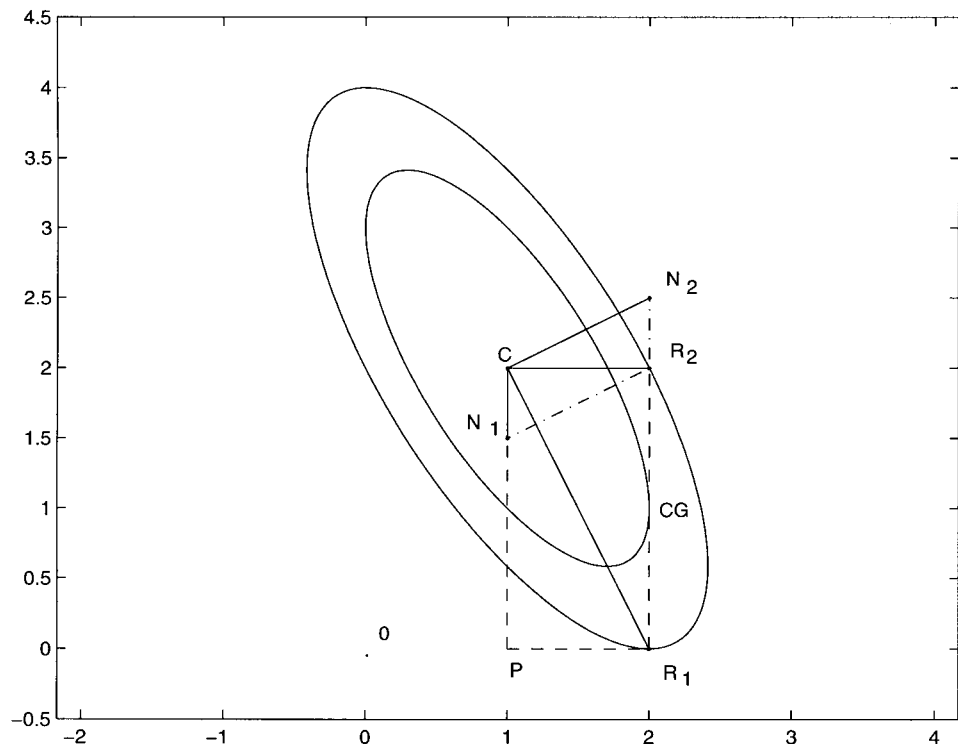
Figure 4. Linear example, initial point (1,3).
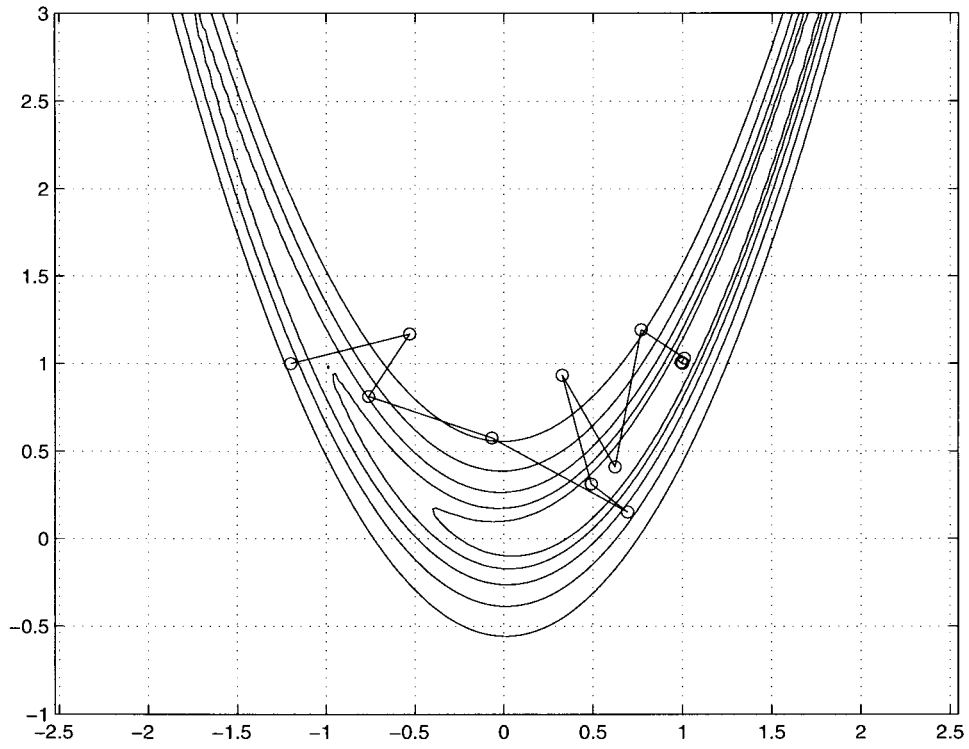

Figure 5. Linear example, initial point (2,0).

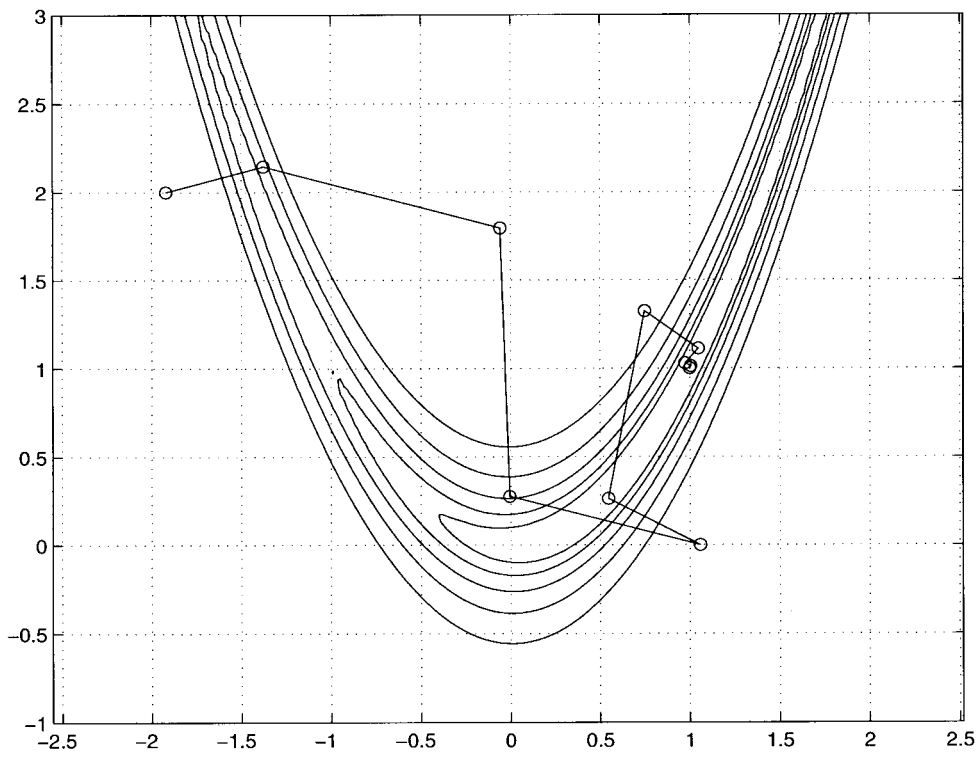*Figure 6. Nonlinear example, initial point (−1.2,1).*

Figure 7. Nonlinear example, initial point (−1.92,2).

12

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE October 1999	3. REPORT TYPE AND DATES COVERED Technical Memorandum

4. TITLE AND SUBTITLE

Method of Conjugate Radii for solving Linear and Nonlinear Systems

5. FUNDING NUMBERS

632-30-34

6. AUTHOR(S)

Philip R. Nachtsheim

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Ames Research Center
Moffett Field, CA 94035-1000

8. PERFORMING ORGANIZATION REPORT NUMBER

A-00V0001

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

NASA/TM–1999-209580

11. SUPPLEMENTARY NOTES

Point of Contact: Philip R. Nachtsheim, Ames Research Center, MS 269-1, Moffett Field, CA 94035-1000
(650) 604-4387

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified — Unlimited
Subject Category 59   Distribution: Standard
Availability: NASA CASI (301) 621-0390

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This paper describes a method to solve a system of N linear equations in N steps. A quadratic form is developed involving the sum of the squares of the residuals of the equations. Equating the quadratic form to a constant yields a surface which is an ellipsoid. For different constants, a family of similar ellipsoids can be generated. Starting at an arbitrary point an orthogonal basis is constructed and the center of the family of similar ellipsoids is found in this basis by a sequence of projections. The coordinates of the center in this basis are the solution of the linear system of equations. A quadratic form in N variables requires N projections. That is, the current method is an exact method. It is shown that the sequence of projections is equivalent to a special case of the Gram-Schmidt orthogonalization process. The current method enjoys an advantage not shared by the classic Method of Conjugate Gradients. The current method can be extended to nonlinear systems without modification. For nonlinear equations the Method of Conjugate Gradients has to be augmented with a line-search procedure. Results for linear and nonlinear problems are presented.

14. SUBJECT TERMS

Document-sharing, access control, user groups, organizational groupware

15. NUMBER OF PAGES

17

16. PRICE CODE

A03

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500